ANALISIS *SURVIVAL* UNTUK DATA TERSENSOR TIPE II MENGGUNAKAN MODEL DISTRIBUSI LOG-LOGISTIK

SKRIPSI

Diajukan kepada Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Negeri Yogyakarta
untuk memenuhi sebagian prasyarat guna
memperoleh gelar Sarjana Sains

Oleh:

DWI RETNO SARI

07305141026

JURUSAN PENDIDIKAN MATEMATIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS NEGERI YOGYAKARTA

2011

PERSETUJUAN

ANALISIS *SURVIVAL* UNTUK DATA TERSENSOR TIPE II MENGGUNAKAN MODEL DISTRIBUSI LOG-LOGISTIK

Oleh:

Dwi Retno Sari

07305141026

Telah disetujui dan disyahkan pada tanggal

27 Juni 2011

untuk diujikan di depan dewan penguji skripsi

Program Studi Matematika

Jurusan Pendidikan Matematika

Fakultas Matematika dan Ilmu Pengetahuan Alam

Universitas Negeri Yogyakarta

Menyetujui,

Pembimbing

Dr. Dhoriva UW

NIP. 196603311993032001

SKRIPSI

ANALISIS *SURVIVAL* UNTUK DATA TERSENSOR TIPE II MENGGUNAKAN MODEL DISTRIBUSI LOG-LOGISTIK

Oleh:
Dwi Retno Sari
07305141026

Telah diujikan di depan dewan penguji skripsi Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Yogyakarta pada tanggal 11 Juli 2011 dan dinyatakan telah memenuhi persyaratan guna memperoleh gelar Sarjana Sains.

DEWAN PENGUJI

Nama	Jabatan	Tanda Tangan	Tanggal
Dr. Dhoriva U.W.	Ketua Penguji	....	27 Juli 2011
Elly Arliani, M.Si.	Sekretaris Penguji	....	26 Juli 2011
Endang L, M.Si.	Penguji Utama	....	18 Juli 2011
M. Susanti, M.Si.	Anggota Penguji	....	25 Juli 2011

Yogyakarta, Juli 2011
Fakultas Matematika dan Ilmu Pengetahuan Alam
Universitas Negeri Yogyakarta
Dekan,

KEMENTERIAN PENDIDIKAN NASIONAL UNIVERSITAS NEGERI YOGYAKARTA FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

Dr. Ariswan
NIP. 19590914 198803 1 003

PERNYATAAN

Yang bertanda tangan di bawah ini:

Nama : Dwi Retno Sari

NIM : 07305141026

Prodi/ Jurusan : Matematika/ Pendidikan Matematika

Fakultas : Matematika dan Ilmu Pengetahuan Alam

Judul TAS : Analisis *Survival* untuk Data Tersensor Tipe II Menggunakan Model Distribusi Log-logistik

Dengan ini saya menyatakan bahwa skripsi ini adalah hasil pekerjaan saya sendiri dan sepanjang pengetahuan saya tidak berisi materi yang dipublikasikan atau ditulis oleh orang lain atau telah digunakan sebagai persyaratan penyelesaian studi di Perguruan Tinggi lain kecuali pada bagian tertentu yang saya ambil sebagai acuan. Apabila ternyata terbukti pernyataan ini tidak benar, sepenuhnya menjadi tanggung jawab saya.

Yogyakarta, Juli 2011

Yang menyatakan,

Dwi Retno Sari
NIM. 07305141026

HALAMAN MOTTO

Aku pasti bisa

(Penulis)

Sesungguhnya sesudah kesulitan itu ada kemudahan

(QS. Alam Nasyrah:6)

Tiadanya keyakinanlah yang membuat orang takut menghadapi tantangan, dan saya percaya pada diri saya sendiri. (Muhammad Ali)

HALAMAN PERSEMBAHAN

Kupersembahkan karya kecil ini untuk:

- *Kedua orangtuaku, Ibu dan Bapak tercinta*

Terima kasih atas doa restu dan kasih sayangnya, sungguh budimu tidak akan bisa terbalaskan.

- *Kakak dan adikku tersayang, Ferita Indriyati dan Adhi Surya H*

Terima kasih atas doa dan dukungannya.

- *Ketiga sahabat terbaikku, Nana, Nurul, Santi*

Terima kasih atas dukungan, motivasi dan semangat dari kalian. kebersamaan kita, tangis dan tawa bersama kalian tak akan pernah aku lupakan.

- *Sahabat-sahabat S.O.V: Anna, Azi, Dhita, Fifi, Ika, Lina, Nawang, Riza, Susi*

Terima kasih semua, atas dukungan, motivasi dan semangat dari kalian. Karna kalian aku bisa menyelesaikan kuliah yang penuh rintangan dengan canda dan tawa bersama kalian.

- *Teman-teman Matematika Reguler 2007*

KATA PENGANTAR

Segala puji dan syukur penulis panjatkan kehadirat Tuhan yang Maha Esa, yang telah memberikan segala rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan skripsi dengan judul "Analisis *Survival* untuk Data Tersensor Tipe II Menggunakan Model Distribusi Log-logistik" ini guna memenuhi persyaratan untuk memperoleh gelar Sarjana Sains pada Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Yogyakarta.

Penulis mengucapkan terima kasih kepada:

1. Bapak Dr. Ariswan, selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam yang telah mendukung penulisan skripsi ini.
2. Bapak Dr. Hartono, selaku Ketua Jurusan Pendidikan Matematika yang telah mendukung penulisan skripsi ini.
3. Ibu Atmini Dhoruri, M.Si, selaku Ketua Program Studi Matematika yang telah mendukung penulisan skripsi ini.
4. Ibu Dr. Dhoriva U.W, selaku dosen pembimbing skripsi yang dengan penuh kesabaran telah meluangkan waktu untuk memberikan bimbingan, saran dan pengarahan dalam menyelesaikan skripsi ini.
5. Seluruh Dosen Jurusan Pendidikan Matematika yang telah memberikan ilmu kepada penulis.
6. Bapak dan Ibu serta keluarga semua yang telah mencurahkan kasih sayang.
7. Teman-teman matematika angkatan 2007 yang telah memberikan bantuan dan dukungan dalam penyelesaian penulisan skripsi ini.

8. Semua pihak yang telah membantu penulis dalam menyelesaikan skripsi ini yang tidak dapat penulis sebutkan satu persatu.

Penulis menyadari bahwa skripsi ini kurang sempurna, semoga menjadi pelajaran bagi para pembaca agar bisa menyempurnakan penulisan selanjutnya. Semoga skripsi ini bermanfaat bagi para pembaca, khususnya para pencinta matematika.

Yogyakarta, Juli 2011

Penulis,

ANALISIS *SURVIVAL* UNTUK DATA TERSENSOR TIPE II MENGGUNAKAN MODEL DISTRIBUSI LOG-LOGISTIK

Oleh:
Dwi Retno Sari
07305141026

ABSTRAK

Analisis *survival* merupakan suatu analisis data mengenai daya tahan hidup atau lamanya waktu hidup suatu individu atau unit pada keadaan tertentu. Tujuan penulisan skripsi ini adalah untuk mendapatkan model *survival* untuk data tersensor tipe II, mendapatkan estimasi parameter-parameter, interval konfidensi untuk parameter-parameter dan contoh penerapannya. Biasanya data *survival* akan mengikuti distribusi tertentu. Dalam skripsi ini akan dibahas mengenai data *survival* yang berdistribusi log-logistik.

Model *survival* untuk data tersensor tipe II ditentukan dengan mencari estimasi parameter-parameter yaitu γ dan β berdasarkan fungsi *maximum likelihood* dan menentukan interval konfidensi untuk tiap-tiap parameter dengan mencari matriks informasi dan matriks kovarian terlebih dahulu. Sedangkan contoh data berdistribusi log-logistik didapatkan dengan metode simulasi pembangkitan data dengan *software Minitab 14.*

Berdasarkan hasil pembahasan maka dapat disimpulkan bahwa model *survival* untuk data tersensor tipe II yang berdistribusi log-logistik yaitu

$$L_L(\gamma,\beta)=\frac{n!}{(n-r)!}\left[\left(\frac{\beta}{\gamma}\right)^r\prod_{i=1}^{r}\frac{\left(\frac{t_i}{\gamma}\right)^{\beta-1}}{\left[1+\left(\frac{t_i}{\gamma}\right)^{\beta}\right]^2}\right]\left[\frac{1}{1+\left(\frac{t_r}{\gamma}\right)^{\beta}}\right]^{n-r}.$$

Estimasi parameter untuk γ dan β pada model *survival* untuk data tersensor tipe II berdasarkan distribusi log-logistik yaitu menggunakan metode *maximum likelihood*, dengan interval konfidensi untuk γ adalah $\hat{\gamma}-z_{\alpha/2}se(\hat{\gamma})\le\gamma\le\hat{\gamma}+z_{\alpha/2}se(\hat{\gamma})$, dan interval konfidensi untuk β adalah $\hat{\beta}-z_{\alpha/2}se(\hat{\beta})\le\beta\le\hat{\beta}+z_{\alpha/2}se(\hat{\beta})$. Dari contoh data umur penyakit pasien penderita kanker paru-paru yang berdistribusi log-logistik, didapatkan estimasi parameter untuk γ dan β adalah 79,61400344 dan 2,14780784. Sedangkan interval konfidensi untuk γ yaitu $56,76518135\le\gamma\le 102,4628255$ dan interval konfidensi untuk β yaitu $1,371336761\le\beta\le 2,924420039$. Peluang hidup untuk pasien yang menderita kanker paru-paru selama 50 bulan adalah 0,73, sedangkan peluang hidup pasien yang menderita selama 90 bulan adalah 0,43.

DAFTAR ISI

Halaman

DAFTAR TABEL

DAFTAR GAMBAR

DAFTAR LAMPIRAN

BAB I

PENDAHULUAN

A. LATAR BELAKANG

Dalam bidang matematika terdapat cabang statistika yang telah berkembang pesat dengan adanya penemuan-penemuan alat analisis yang dapat digunakan untuk menganalisis suatu permasalahan. Salah satunya adalah uji hidup yang merupakan penelitian daya tahan hidup suatu unit atau individu pada suatu keadaan tertentu. Uji hidup biasa digunakan dalam bidang teknik, biologi, kedokteran dan lain-lain. Penelitian-penelitian tersebut biasanya menggunakan data yang berkaitan dengan waktu hidup dari suatu individu. Analisis yang digunakan untuk menganalisis data waktu hidup tersebut disebut analisis *survival*. Analisis *survival* mencakup berbagai teknik statistik yang berguna untuk menganalisis berbagai macam variabel random positif. Variabel random positif pada analisis *survival* berupa *survival time* (waktu tahan hidup) atau *failure time* (waktu kegagalan).

Dalam penelitian uji hidup, data waktu hidup dapat berbentuk data lengkap, data tersensor tipe I dan data tersensor tipe II. Data tersebut lengkap jika data diamati secara utuh. Data tersensor tipe I merupakan data uji hidup yang dihasilkan setelah penelitian berjalan selama waktu yang telah ditentukan. Sedangkan data tersensor tipe II merupakan data hasil penelitian dimana penelitian dihentikan setelah kematian atau kegagalan tertentu telah terjadi (Lawless, 1982).

Data tersensor tipe II adalah suatu data waktu hidup yang terdapat r buah observasi dalam sampel random yang berukuran *n* dengan $1 \leq r \leq n$. Dalam suatu penelitian, penyensoran tipe II lebih sering digunakan, karena dalam uji hidup ini terdapat observasi sebanyak *n*, tetapi penelitian dihentikan ketika observasi mengalami kegagalan ke-*r*, sehingga peneliti dapat menghemat waktu dan biaya.

Untuk menganalisis data *survival* dengan data tersensor diperlukan asumsi tertentu tentang distribusi populasinya. Beberapa distribusi parametrik yang populer dan dapat digunakan untuk menganalisis model *survival* adalah Distribusi Weibull, Distribusi Eksponensial, Distribusi Log-normal, Distribusi Gamma, Distribusi Log-logistik dan lain-lain.

Dari beberapa distribusi yang ada, skripsi ini menggunakan fungsi *survival* berdistribusi Log-logistik, atau data waktu hidup diasumsikan mengikuti Distribusi Log-logistik. Distribusi Log-logistik masih jarang digunakan dalam analisis *survival*. Distribusi Log-logistik mempunyai bentuk yang hampir sama dengan Distribusi Log-normal. Misalnya untuk meneliti tahan hidup pasien yang terserang penyakit kronis, selain itu di bidang industri, juga untuk analisis tahan hidup komponen dari suatu produk. Oleh karena itu penulis mengangkat judul "Analisis *Survival* Untuk Data Tersensor Tipe II Menggunakan Model Distribusi Log-logistik", untuk menentukan analisis *survival* untuk data tersensor tipe II.

B. RUMUSAN MASALAH

Berdasarkan uraian latar belakang, maka rumusan masalah yang akan dibahas adalah:

1. Bagaimana model *survival* untuk data tersensor tipe II berdasarkan model distribusi log-logistik?
2. Bagaimana estimasi parameter model *survival* untuk data tersensor tipe II berdasarkan model distribusi log-logistik?
3. Bagaimana penerapan model *survival* untuk data tersensor tipe II berdasarkan model distribusi log-logistik?

C. TUJUAN PENULISAN

Tujuan dalam penulisan skripsi ini adalah:

1. Mendapatkan model *survival* untuk data tersensor tipe II berdasarkan model distribusi log-logistik.
2. Mendapatkan estimasi parameter model *survival* untuk data tersensor tipe II berdasarkan model distribusi log-logistik.
3. Menjelaskan penerapan model *survival* untuk data tersensor tipe II berdasarkan model distribusi log-logistik

D. MANFAAT PENULISAN

Manfaat dari penulisan ini adalah :

1. Menambah referensi tentang analisis *survival*, khususnya data tersensor tipe II menggunakan distribusi log-logistik.
2. Menambah pengetahuan tentang penerapan analisis *survival* untuk data tersensor tipe II menggunakan distribusi log-logistik.

BAB II

LANDASAN TEORI

A. Konsep Dasar Peluang

Pada dasarnya statistika berkaitan dengan penyajian dan penafsiran hasil yang berkemungkinan (hasil yang belum dapat ditentukan sebelumnya) yang muncul dalam penelitian yang dirancang sebelumnya atau yang muncul dalam penelitian ilmiah. Para statistisi berurusan dengan pencacahan atau pengukuran karakteristik suatu objek kajian yang hasilnya berbentuk bilangan. Pekerjaan seperti ini biasa disebut percobaan acak (Abadyo dan Hendro permadi, 2005).

Himpunan semua hasil yang mungkin dari suatu percobaan acak disebut ruang sampel dan dinyatakan dengan lambang **S**. Suatu kejadian adalah himpunan bagian dari ruang sampel (Bain dan Engelhardt, 1992).

Ruang nol atau ruang kosong atau himpunan kosong ialah himpunan bagian ruang sampel yang tidak mengandung satu pun anggota. Kejadian seperti ini dinyatakan dengan lambang ∅ (Walpole, 1995).

Menurut Bain dan Engelhardt (1992), untuk sebuah percobaan, S merupakan ruang sampel dan $A, A_1, A_2, ...$ merepresentasikan kejadian-kejadian yang mungkin. Himpunan fungsi yang menghubungkan nilai P(A) dengan setiap kejadian A disebut himpunan fungsi peluang, dan P(A) merupakan peluang dari A, jika memenuhi keadaan sebagai berikut:

1. $0 \leq P(A)$ untuk setiap A
2. $P(S) = 1$

3. $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i)$

dan jika $A_1, A_2, ...$ merupakan kejadian-kejadian yang saling asing.

B. Variabel Random

Variabel random X adalah suatu fungsi dengan daerah asal S, dimana S adalah suatu ruang sampel dan daerah hasil bilangan real sedemikian sehingga *X(e)* = *x*, dengan $e \in S$ dan $x \in \Re$ (Bain dan Engelhardt, 1992).

Terdapat dua macam variabel random, yaitu variabel random diskret dan variabel random kontinu. Jika semua harga yang mungkin dari variabel random *X*, adalah himpunan terhitung (*countable*), $\{x_1, x_2, ..., x_n\}$ atau $\{x_1, x_2, ...\}$, maka X disebut variabel random diskret.

Fungsi *f(x)* = *P*[*X* = *x*] dengan $x = x_1, x_2, ...$ yang memberikan nilai peluang untuk setiap *X* yang mungkin, disebut fungsi densitas peluang diskret *(discrete probability density function / discrete pdf)*.

Fungsi $F(x) = P[X \le x]$ merupakan fungsi distribusi kumulatif (*cumulative distribution function / CDF*) dari variabel random *X* untuk sembarang bilangan real *x*.

Menurut Bain dan Engelhardt (1992), jika himpunan semua nilai yang mungkin dari suatu variabel random *X* merupakan selang bilangan real, maka *X* disebut variabel random kontinu.

Suatu fungsi *f(x)* yang didefinisikan pada selang nilai variabel random *X* disebut fungsi densitas peluang, sehingga fungsi distribusi kumulatifnya dapat dinyatakan sebagai

$$F(x) = \int_{-\infty}^{x} f(t)dt \, .$$

C. Konsep Dasar Distribusi *Survival*

Data *survival* adalah data lamanya individu-individu atau unit-unit dari suatu populasi menjalankan fungsinya dengan baik sampai kematian individu-individu tersebut. Dalam mempelajari penerapan data *survival*, terlebih dahulu harus diketahui konsep-konsep statistik pada distribusi *survival*.

Misalkan T merupakan variabel random kontinu non negatif yang menunjukkan tahan hidup individu-individu dari suatu populasi. Pada model kontinu, fungsi-fungsi seperti fungsi densitas peluang, fungsi distribusi kumulatif, fungsi *hazard* dan fungsi *survivor* didefinisikan dalam interval $[0, \infty)$ (Lawless, 1982).

Fungsi densitas peluang pada analisis *survival* adalah peluang suatu individu mati atau gagal dalam interval waktu *t* sampai $t + \Delta t$, dengan waktu T merupakan variabel random. Fungsi densitas peluang dari T dapat dinyatakan sebagai *f(t)*,

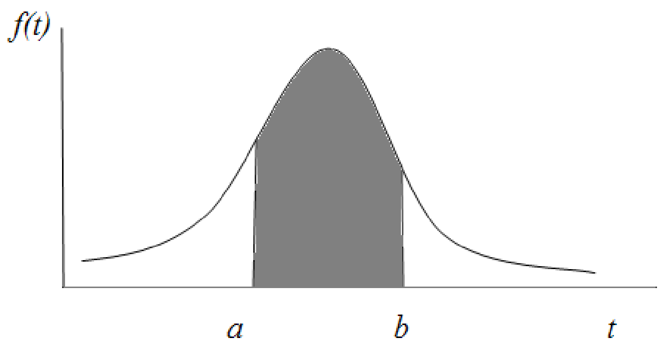
$$f(t) = \lim_{\Delta t \to 0} \left[\frac{P(t \le T < (t + \Delta t))}{\Delta t} \right]$$

yang mempunyai sifat sebagai berikut:

a. $f(t) \ge 0 \, , \; t \ge 0$

b. $\int_0^\infty f(t)dt = 1$

Fungsi *f* disebut fungsi densitas peluang bagi variabel random kontinu T bila luas daerah di bawah kurva dan di atas sumbu-*t* sama dengan 1, dan bila luas daerah di bawah kurva antara *t*=*a* dan *t*=*b* menyatakan peluang T terletak antara *a* dan *b* (Walpole, 1995), sebagaimana diilustrasikan dalam gambar 2.1.



Gambar 2.1 Kurva fungsi densitas peluang

Dengan demikian luas daerah yang diarsir adalah $P(a < T < b) = \int_a^b f(t)dt$ dengan $a, b \in [0, \infty)$.

1. Fungsi Distribusi Kumulatif

Jika T merupakan variabel random dari waktu hidup suatu individu dalam interval $[0, \infty)$, maka fungsi distribusi kumulatif *F(t)* untuk distribusi kontinu dengan fungsi densitas peluang *f(t)* dinyatakan sebagi berikut (Lawless, 1982):

$$F(t) = P(T \le t)$$

Atau

$$F(t) = \int_0^t f(x)dx \text{, untuk } t > 0$$

2. Fungsi *Survivor*

Menurut Lawless (1982) fungsi *survivor* didefinisikan sebagai peluang suatu individu dapat bertahan hidup sampai waktu *t*. Jika *T* merupakan variabel random dari waktu hidup suatu individu dalam interval $[0,\infty)$, maka fungsi *survivor* S(t) dapat dinyatakan dalam persamaan:

$$S(t) = P(T \geq t)$$

$$= \int_t^\infty f(x)dx$$

Dengan demikian diperoleh persamaan yang menyatakan hubungan antara fungsi *survivor* dan fungsi distribusi kumulatif, yaitu

$$S(t) = 1 - F(t)$$

3. Fungsi *Hazard*

Fungsi *hazard* menyatakan peluang kegagalan suatu komponen pada waktu *t*, jika diketahui bahwa komponen tersebut tetap hidup hingga waktu *t*. Menurut Lawless (1982) fungsi *hazard* adalah peluang suatu individu mati dalam interval waktu *t* sampai $t + \Delta t$, jika diketahui individu tersebut masih dapat bertahan hidup sampai dengan waktu *t*, yang dinyatakan sebagai berikut:

$$h(t) = \lim_{\Delta t \to 0} \left[\frac{P(t \leq T < t + \Delta t \,|\, T \geq t)}{\Delta t} \right]$$

Jika f(t) adalah fungsi densitas peluang pada waktu *t*, maka diperoleh

$$h(t) \quad = \lim_{\Delta t \to 0} \left[\frac{P(t \leq T < t + \Delta t \,|\, T \geq t)}{\Delta t} \right]$$

$$= \lim_{\Delta t \to 0} \left[\frac{P[(t \leq T < (t + \Delta t)) \cap (T \geq t)]}{P(T \geq t).\Delta t} \right]$$

$$= \lim_{\Delta t \to 0} \left[\frac{P(t \le T < (t + \Delta t))}{P(T \ge t).\Delta t} \right]$$

$$= \lim_{\Delta t \to 0} \left[\frac{1}{\Delta t} . \frac{F(t + \Delta t) - F(t)}{1 - F(t)} \right]$$

$$= \lim_{\Delta t \to 0} \left[\frac{F(t + \Delta t) - F(t)}{\Delta t} . \frac{1}{S(t)} \right]$$

$$= \frac{F'(t)}{S(t)}$$

$$h(t) = \frac{f(t)}{S(t)}.$$

D. Data Tersensor

Dalam penelitian uji hidup, data waktu hidup dapat berbentuk data lengkap, data tersensor tipe I dan data tersensor tipe II. Pada pengambilan data menggunakan data lengkap, percobaan akan dihentikan jika semua komponen atau individu yang diteliti gagal atau mati (Lawless, 1982). Metode menggunakan data lengkap memerlukan waktu yang lama sehingga jarang digunakan.

Data tersensor adalah data yang diperoleh sebelum semua data teramati waktu hidupnya, sedangkan waktu pengamatan telah berakhir atau oleh sebab lain. Data tersensor tipe I merupakan data uji hidup yang dihasilkan setelah penelitian berjalan selama waktu yang telah ditentukan. Sedangkan data tersensor tipe II merupakan data hasil penelitian dimana penelitian dihentikan setelah kematian atau kegagalan tertentu telah terjadi (Lawless, 1982).

Data tersensor tipe II merupakan data kematian atau kegagalan yang tidak lengkap (*incomplete mortality data*) yaitu data waktu kematian atau kegagalan dari r observasi terkecil dalam sampel random yang berukuran n dengan $1 \leq r \leq n$. Dalam suatu penelitian, penyensoran tipe II lebih sering digunakan, yaitu dalam uji hidup yang terdapat observasi sebanyak n, tetapi penelitian dihentikan ketika observasi mengalami kegagalan ke-r, sehingga dapat menghemat waktu dan biaya. Dalam penyensoran ini, r ditentukan terlebih dahulu sebelum data dikumpulkan.

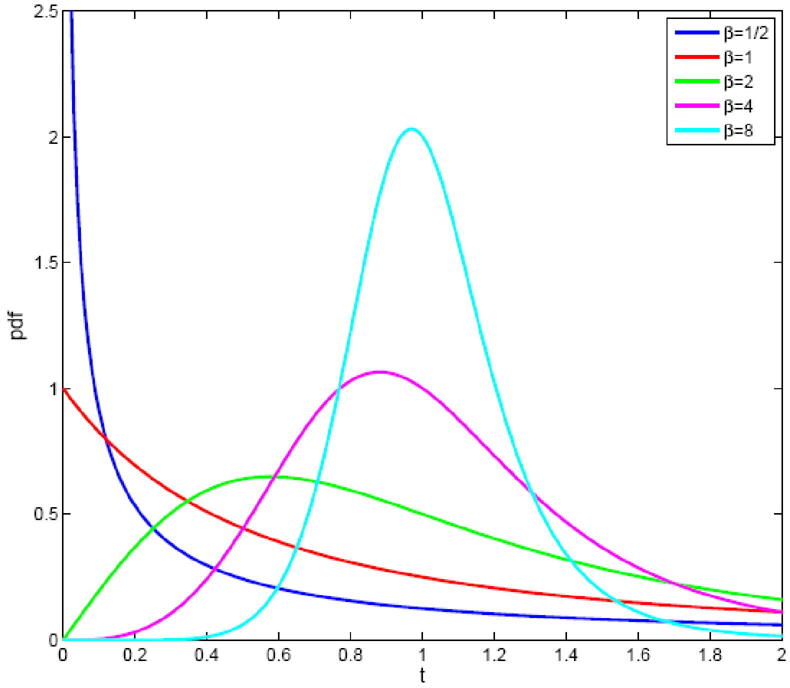
E. Distribusi Log-logistik

Dalam statistika, distribusi log-logistik merupakan salah satu distribusi peluang kontinu untuk variabel random non-negatif. Distribusi ini digunakan dalam analisis tahan hidup sebagai model parametrik, misalnya untuk meneliti waktu penyembuhan suatu penyakit. Distribusi log-logistik juga telah dikembangkan di bidang industri untuk menganalisis tahan hidup komponen dari suatu produk.

Variabel random T dikatakan mengikuti distribusi log-logistik dengan parameter γ dan parameter *shape* β, jika mempunyai fungsi densitas:

$$f(t;\gamma,\beta)=\frac{\left(\frac{\beta}{\gamma}\right)\left(\frac{t}{\gamma}\right)^{\beta-1}}{\left[1+\left(\frac{t}{\gamma}\right)^{\beta}\right]^{2}}\ ,\ t>0\ ,\ \text{dimana}\ \gamma>0\ \text{dan}\ \beta>0\ .$$

untuk selanjutnya dinotasikan sebagai $T \sim L_L(\gamma,\beta)$. Nilai parameter *shape* yaitu β menyatakan suatu bentuk yang bermacam-macam dari kurva fungsi densitas yaitu naik, turun, atau mendatar, sehingga kondisi ini sangat cocok digunakan untuk berbagai model data *survival*.



Gambar 2.2 Kurva fungsi densitas peluang dari distribusi log-logistik

Fungsi densitas peluang dari distribusi log-logistik dengan $\gamma = 1$ ditunjukkan pada gambar 2.2 untuk nilai β yang berbeda. Untuk $0 \le \beta \le 1$ fungsi densitas peluangnya menurun, sedangkan untuk $\beta > 1$ fungsi densitas peluangnya merupakan fungsi naik dengan sebuah puncak. Semakin besar nilai β, puncak dari kurva fungsi densitas peluangnya semakin runcing dan bentuknya semakin simetris.

Fungsi distribusi kumulatifnya adalah

$$F(t) = \int_0^t \frac{\left(\frac{\beta}{\gamma}\right)\left(\frac{y}{\gamma}\right)^{\beta-1}}{\left[1+\left(\frac{y}{\gamma}\right)^{\beta}\right]^2} dy$$

misal: $u = 1 + \left(\frac{y}{\gamma}\right)^{\beta}$

$$du = \left(\frac{\beta}{\gamma}\right)\left(\frac{y}{\gamma}\right)^{\beta-1} dy$$

$$dy = \frac{du}{\left(\frac{\beta}{\gamma}\right)\left(\frac{y}{\gamma}\right)^{\beta-1}}$$

$$F(t) = \int_0^t \frac{\left(\frac{\beta}{\gamma}\right)\left(\frac{y}{\gamma}\right)^{\beta-1}}{\left[1+\left(\frac{y}{\gamma}\right)^{\beta}\right]^2} dy$$

$$= \int_0^t \frac{\left(\frac{\beta}{\gamma}\right)\left(\frac{y}{\gamma}\right)^{\beta-1}}{u^2} . \frac{du}{\left(\frac{\beta}{\gamma}\right)\left(\frac{y}{\gamma}\right)^{\beta-1}}$$

$$= \int_0^t u^{-2} du$$

$$= -\left(1+\left(\frac{y}{\gamma}\right)^{\beta}\right)^{-1}\Bigg|_0^t$$

$$= -\frac{1}{1+\left(\frac{t}{\gamma}\right)^{\beta}} + 1$$

$$= 1 - \frac{1}{1+\left(\frac{t}{\gamma}\right)^{\beta}}$$

$$= \frac{1+\left(\frac{t}{\gamma}\right)^{\beta} - 1}{1+\left(\frac{t}{\gamma}\right)^{\beta}}$$

$$= \frac{\left(\frac{t}{\gamma}\right)^{\beta}}{1+\left(\frac{t}{\gamma}\right)^{\beta}}$$

$$= \frac{t^{\beta}}{\gamma^{\beta}+t^{\beta}}$$

Jadi $F(t;\gamma,\beta)=\dfrac{t^{\beta}}{\gamma^{\beta}+t^{\beta}}$, $t>0$, dimana $\gamma>0$ dan $\beta>0$

Fungsi *survivor* dari $T\sim L_L(\gamma,\beta)$ didefinisikan sebagai peluang suatu individu dapat bertahan hidup sampai waktu *t*, yaitu

$$
\begin{aligned}
S(t) &= 1-F(t)\\
&= 1-\frac{t^{\beta}}{\gamma^{\beta}+t^{\beta}}\\
&= \frac{\gamma^{\beta}}{\gamma^{\beta}+t^{\beta}}\\
&= \frac{1}{1+\left(\frac{t}{\gamma}\right)^{\beta}}
\end{aligned}
$$

Fungsi *hazard* *h(t)* menyatakan peluang suatu komponen mengalami kegagalan pada waktu *t*.

$$
\begin{aligned}
h(t) &= \frac{f(t)}{S(t)}\\
&= \frac{\left(\frac{\beta}{\gamma}\right)\left(\frac{t}{\gamma}\right)^{\beta-1}}{\left[1+\left(\frac{t}{\gamma}\right)^{\beta}\right]}
\end{aligned}
$$

F. Metode Maksimum Likelihood

Metode maksimum Likelihood adalah salah satu metode yang paling sering digunakan untuk mencari nilai estimasi dari suatu parameter. Fungsi kepadatan bersama *(joint density function)* dari n variabel random $X_1, X_2, \ldots, X_n$ pada $x_1, x_2, \ldots, x_n$ adalah $f(x_1, x_2, \ldots, x_n;\ \theta)$ disebut sebagai fungsi likelihood.

Untuk $x_1, x_2, ..., x_n$ yang tetap, fungsi likelihood adalah fungsi dari θ dan sering dinotasikan sebagai L(θ).

Jika $X_1, X_2, ..., X_n$ menyatakan sampel random dengan fungsi densitas peluang f(x; θ) maka:

$$L(\theta) = f(x_1;\theta)...f(x_n;\theta) = \prod_{i=1}^{n} f(x_i;\theta)$$

Misalkan $L(\theta) = f(x_1;\theta)...f(x_n;\theta)$, $\theta \in \Omega$ adalah fungsi kepadatan bersama dari $X_1, X_2, ..., X_n$. Untuk sekumpulan observasi yang diberikan $(x_1, x_2, ..., x_n)$, suatu nilai $\hat{\theta}$ dalam Ω sedemikian hingga $L(\theta)$ maksimum, disebut *Maximum Likelihood Estimator (MLE)* dari θ. Nilai $\hat{\theta}$ adalah nilai θ yang memenuhi:

$$f(x_1, x_2, ..., x_n;\hat{\theta}) = \max_{\theta \in \Omega} f(x_1, x_2, ..., x_n;\theta)$$

Apabila Ω adalah interval terbuka, dan jika $L(\theta)$ adalah differensiabel dan diasumsikan maksimum pada Ω maka MLE adalah solusi dari persamaan:

$$\frac{d}{d\theta} L(\theta) = 0$$

Hal yang perlu diperhatikan, jika ternyata terdapat lebih dari satu solusi untuk persamaan $\frac{d}{d\theta} L(\theta) = 0$, maka harus dilakukan perhitungan terhadap masing-masing solusi untuk memperoleh solusi yang memaksimumkan $L(\theta)$. Hal ini dilakukan dengan mencari nilai turunan kedua dari $L(\theta)$, bila nilainya negatif maka solusi tersebut adalah solusi yang maksimum.

Definisi tentang fungsi likelihood dan estimasi kemungkinan maksimum dapat diterapkan dalam parameter-parameter tak diketahui yang lebih dari satu. Bila **θ** adalah parameter, katakan $\theta = (\theta_1, \theta_2, ..., , \theta_k)$, maka estimasi kemungkinan maksimumnya akan berupa persamaan simultan dengan penurunan parsial tiap-tiap parameternya.

$$\frac{\partial}{\partial \theta_j} \ln L(\theta_1, \theta_2, ..., \theta_k) = 0$$

untuk j =1, 2, ... ,k.

Persamaan di atas disebut persamaan kemungkinan maksimum (*Maximum Likelihood Equations*). Nilai $\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_k$ merupakan estimator bila persamaan kemungkinan maksimumnya memberikan nilai maksimum terhadap $L(\theta_1, \theta_2, ..., \theta_k)$.

G. Statistik Terurut

Misalkan himpunan variabel random $X_1, X_2, ..., X_n$ merupakan sampel random yang berukuran n dari suatu populasi dengan fungsi densitas *f(x)* maka fungsi densitas peluang bersama dari variabel random independennya adalah sebagai berikut (Bain dan Engelhardt, 1992):

$$f(x_1, x_2, ..., x_n) = f(x_1) f(x_2) ... f(x_n)$$

Misalkan $X_1, X_2, ..., X_n$ adalah sampel random yang berukuran n dari fungsi densitas peluang *f(x)*, dimana untuk *f(x)* kontinu dan *f(x)* > 0, $a < x < b$, maka fungsi densitas peluang dari statistik terurut ke-*k*, Y_k adalah

$$g_k(y_k) = \frac{n!}{(k-1)!(n-k)!} \left[F(y_k)\right]^{k-1} \left[1 - F(y_k)\right]^{n-k} f(y_k) \quad \text{, jika } a < y_k < b.$$

H. Matriks Informasi

Misalkan $y_1, y_2, ..., y_n$ merupakan sampel random dari suatu distribusi dengan fungsi densitas peluang *f(y;**θ**)*, dimana $\boldsymbol{\theta} = (\theta_1, ..., \theta_k)'$ merupakan vektor dari parameter-parameter yang belum diketahui nilainya yang merupakan himpunan bagian dari Ω. Fungsi likelihood dari θ adalah

$$L(\theta) = \prod_{i=1}^{n} f(y_i; \theta)$$

sehingga persamaan maximum likelihoodnya adalah

$$U_j(\theta) = \frac{\partial l(\theta)}{\partial \theta_j} \ , \ j = 1, ..., k \, .$$

Menurut Lawless (2003), $U(\theta)$ mempunyai rata-rata **0** dan matriks kovarian $I(\theta)^{-1}$, dimana $I_{ij}(\theta) = E\left(\frac{-\partial^2 \log L(\theta)}{\partial \theta_i \partial \theta_j} \right)$, $i, j = 1, ..., k$. sehingga matrik $I(\theta)$ disebut matriks informasi.

I. Interval Konfidensi

Misalkan $X_1, X_2 ..., X_n$ mempunyai fungsi densitas peluang bersama $f(x_1, x_2 ..., x_n; \theta); \theta \in \Omega$, dimana Ω suatu interval dan misalkan $L = \ell(X_1, X_2 ..., X_n)$ dan $U = a(X_1, X_2 ..., X_n)$. Suatu interval $\left(\ell(x_1, x_2 ..., x_n), a(x_1, x_2 ..., x_n)\right)$ merupakan interval konfidensi $100\alpha\%$ untuk θ jika

$$P\left[\ell(X_1, X_2 ..., X_n) < \theta < a(X_1, X_2 ..., X_n)\right] = \alpha$$

dimana $0 < \alpha < 1$. Sedangkan nilai dari $\ell(x_1, x_2 ..., x_n)$ dan $a(x_1, x_2 ..., x_n)$ disebut batas konfidensi bawah dan batas konfidensi atas (Bain dan Engelhardt, 1992).

Jika $Q = q(X_1, X_2, ..., X_n; \theta)$ adalah sebuah variabel random dari fungsi $X_1, X_2, ..., X_n$ dan θ, maka Q disebut nilai pivot, jika distribusinya tidak tergantung pada θ atau parameter-parameter lain yang tidak diketahui (Bain dan Engelhardt, 1992).

Estimasi interval konfidensi untuk θ dapat diperoleh dengan menggunakan $(\hat{\theta})$ sebagai normal bivariat dengan rata-rata (θ) dan matrik kovarian $\hat{V} = I(\theta)^{-1}$, sehingga standard error untuk θ adalah $se(\hat{\theta}) = \widehat{As\,\mathrm{var}}(\hat{\theta})^{1/2}$ dimana *Asvar* merupakan variansinya (Lawless, 2003).

Interval konfidensi untuk suatu fungsi parametrik $\psi = g(\theta)$ menggunakan pendekatan normal $\hat{\psi} \sim N(\psi, \hat{V}_\psi)$. Dengan demikian nilai pivot dengan pendekatan normalnya adalah

$$Z = \frac{\hat{\psi} - \psi}{\hat{V}_\psi^{1/2}}$$

dan interval konfidensi $1-\alpha$ untuk ψ adalah $\hat{\psi} \pm z_{\alpha/2} \hat{V}_\psi^{1/2}$ (Lawless, 2003).

BAB III

PEMBAHASAN

Dalam pembahasan ini akan dijelaskan mengenai model *survival* dari data tersensor tipe II berdasarkan model distribusi log-logistik. Model *survival* ini ditentukan berdasarkan fungsi *likelihood*nya. Selanjutnya akan dibahas mengenai penerapan model survival untuk data tersensor tipe II berdasarkan model distribusi log-logistik.

A. Data Tersensor Tipe II

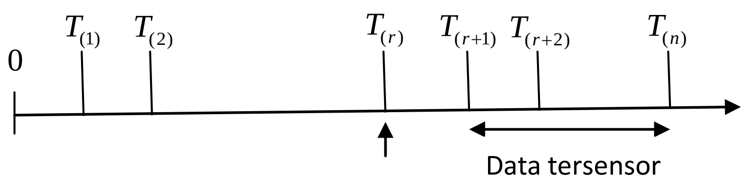
Data tersensor tipe II adalah suatu data waktu kematian atau waktu tahan hidup yang hanya terdapat *r* buah observasi dalam sampel random yang berukuran n dengan $1 \leq r \leq n$. Kebanyakan penelitian menunjukkan penyensoran tipe II lebih sering digunakan, karena dapat menghemat waktu dan biaya. Dalam uji hidup ini, total observasi sebanyak *n*, tetapi uji akan berhenti pada waktu observasi sampel mempunyai waktu kematian atau kegagalan ke-r untuk $1 \leq r \leq n$ (Lawless, 1982).

B. Model *Survival* Data Tersensor Tipe II

Menurut Lawless (2003), misalkan T merupakan variabel random tahan hidup dengan T adalah variabel random kontinu non negatif yang menunjukkan tahan hidup individu-individu dalam suatu populasi yang berdistribusi Log-logistik dengan fungsi densitas peluang yaitu $f(t;\gamma,\beta) = \dfrac{\left(\dfrac{\beta}{\gamma}\right)\left(\dfrac{t}{\gamma}\right)^{\beta-1}}{\left[1+\left(\dfrac{t}{\gamma}\right)^{\beta}\right]^2}$, $t > 0$,

$\gamma > 0$ dan $\beta > 0$. Fungsi *survivor*nya adalah $S(t;\gamma,\beta) = \left[1+\left(\frac{t}{\gamma}\right)^{\beta}\right]^{-1}$ dan fungsi *hazard*nya adalah $h(t;\gamma,\beta) = \frac{\left(\frac{\beta}{\gamma}\right)\left(\frac{t}{\gamma}\right)^{\beta-1}}{\left[1+\left(\frac{t}{\gamma}\right)^{\beta}\right]}$.

Dalam data tersensor tipe II, terdapat r pengamatan dari n sampel yang diamati, dan eksperimen akan dihentikan setelah kegagalan ke-r yang terjadi sebelum waktu t_i. Data terdiri dari *r* tahan hidup terkecil $T_{(1)} \le T_{(2)} \le T_{(3)} \le \ldots \le T_{(r)}$ dari sampel random yang terdiri dari *n* tahan hidup $T_1, T_2, T_3, \ldots, T_n$, seperi diilustrasikan pada Gambar 3.1.



Gambar 3.1. Ilustrasi model tersensor tipe II

Misalkan *T* merupakan variabel random dari *n* individu yang diamati, $f(t_1)$ merupakan fungsi densitas peluang dari variabel random individu ke-1, $f(t_2)$ merupakan fungsi densitas peluang dari variabel random individu ke-2, dan seterusnya hingga $f(t_r)$ untuk variabel random individu ke-*r*.

Individu yang gagal, yaitu individu ke-1 sampai individu ke-r masing-masing sebanyak satu komponen. Sedangkan individu yang masih bertahan melebihi kegagalan dari individu ke-r dituliskan dengan $T_{r+1}, T_{r+2}, T_{r+3}, ..., T_n$ sebanyak *n-r*. Sampel random berukuran n dengan kegagalan r ini mengikuti distribusi multinomial, sehingga terdapat $\frac{n!}{1!1!...1!(n-r)!}$ urutan yang mungkin terjadi dari n pengamatan.

Fungsi densitas peluang bersama dari $T_1, T_2, T_3, ..., T_n$ dari data yang diamati dapat ditulis sebagai berikut:

$$
\begin{aligned}
f(t_1, t_2, ..., t_r) &= \frac{n!}{(n-r)!} f(t_1) f(t_2) ... f(t_r) \left[P(T_{r+1} \geq t_r) ... P(T_n \geq t_r) \right] \\
&= \frac{n!}{(n-r)!} \left[\prod_{i=1}^{r} f(t_i) \right] \left[\left(1 - P\left(T_{r+1} < t_r\right)\right) ... \left(1 - P\left(T_n < t_r\right)\right) \right] \\
&= \frac{n!}{(n-r)!} \left[\prod_{i=1}^{r} f(t_i) \right] \left[\left(1 - F(t_r)\right) ... \left(1 - F(t_r)\right) \right] \\
&= \frac{n!}{(n-r)!} \left[\prod_{i=1}^{r} f(t_i) \right] \left[1 - F(t_r) \right]^{n-r} \\
&= \frac{n!}{(n-r)!} \left[\prod_{i=1}^{r} f(t_i) \right] \left[S(t_r) \right]^{n-r}
\end{aligned}
$$

Fungsi densitas peluang bersama data tersensor tipe II dari $t_1, t_2, ..., t_r$ untuk $r < n$ adalah $f(t_1, t_2, ..., t_r) = \frac{n!}{(n-r)!} \left[\prod_{i=1}^{r} f(t_i) \right] [S(t_r)]^{n-r}$. Karena diketahui bahwa

$f(t_i) = \frac{\left(\frac{\beta}{\gamma}\right)\left(\frac{t_i}{\gamma}\right)^{\beta-1}}{\left[1 + \left(\frac{t_i}{\gamma}\right)^{\beta}\right]^2}$ dan $S(t_r) = \left[1 + \left(\frac{t_r}{\gamma}\right)^{\beta}\right]^{-1}$, maka fungsi *likelihood*nya adalah sebagai berikut

$$f(t_1,t_2,...,t_r)=\frac{n!}{(n-r)!}\left[\prod_{i=1}^{r} f(t_i)\right][S(t_r)]^{n-r} \tag{3.1}$$

$$=\frac{n!}{(n-r)!}\left[\prod_{i=1}^{r}\frac{\left(\frac{\beta}{\gamma}\right)\left(\frac{t_i}{\gamma}\right)^{\beta-1}}{\left[1+\left(\frac{t_i}{\gamma}\right)^{\beta}\right]^2}\right]\left[\frac{1}{1+\left(\frac{t_r}{\gamma}\right)^{\beta}}\right]^{n-r}$$

$$=\frac{n!}{(n-r)!}\left[\left(\frac{\beta}{\gamma}\right)^r\prod_{i=1}^{r}\frac{\left(\frac{t_i}{\gamma}\right)^{\beta-1}}{\left[1+\left(\frac{t_i}{\gamma}\right)^{\beta}\right]^2}\right]\left[\frac{1}{1+\left(\frac{t_r}{\gamma}\right)^{\beta}}\right]^{n-r}$$

Jadi fungsi *likelihood* dari distribusi log-logistik untuk data tersensor tipe II memiliki bentuk

$$L_L(\gamma,\beta)=\frac{n!}{(n-r)!}\left[\left(\frac{\beta}{\gamma}\right)^r\prod_{i=1}^{r}\frac{\left(\frac{t_i}{\gamma}\right)^{\beta-1}}{\left[1+\left(\frac{t_i}{\gamma}\right)^{\beta}\right]^2}\right]\left[\frac{1}{1+\left(\frac{t_r}{\gamma}\right)^{\beta}}\right]^{n-r} \tag{3.2}$$

C. *Maximum Likelihood Estimator (MLE)*

Dalam analisis data *survival* terlebih dahulu dipilih bentuk distribusi dari data, kemudian dicari bentuk fungsi parameter yang diwakili data survival tersebut. Dalam skripsi ini menggunakan metode maksimum *likelihood* untuk mencari estimasi parameter dari distribusi log-logistik. Metode maksimum *likelihood* menggunakan nilai dalam ruang parameter Ω yang bersesuaian dengan harga kemungkinan maksimum dari data observasi sebagai estimasi dari parameter yang tidak diketahui.

Dalam aplikasinya $L_L(\gamma,\beta)$ menunjukkan fungsi densitas peluang bersama dari sampel random. Jika Ω ruang parameter yang merupakan interval terbuka dan

$L_L(\gamma,\beta)$ merupakan fungsi yang dapat diturunkan serta diasumsikan maksimum pada Ω, maka persamaan maksimum *likelihood*nya adalah

$$\frac{\partial L_L(\gamma,\beta)}{\partial\gamma}=0 \text{ dan } \frac{\partial L_L(\gamma,\beta)}{\partial\beta}=0$$

Jika penyelesaian dari persamaan tersebut ada, maka maksimum dari $L_L(\gamma,\beta)$ dapat terpenuhi. Apabila penyelesaian dari persamaan tersebut sulit untuk diselesaikan maka fungsi $L_L(\gamma,\beta)$ dapat dibuat logaritma naturalnya, dengan ketentuan $\ln L_L(\gamma,\beta)$ maksimum, sehingga persamaan logaritma natural maksimum *likelihood*nya adalah

$$\frac{\partial \ln L_L(\gamma,\beta)}{\partial\gamma}=0 \text{ dan } \frac{\partial \ln L_L(\gamma,\beta)}{\partial\beta}=0$$

Untuk mengetahui apakah penduga dari γ dan β tersebut telah maksimum, maka dicari turunan ke-2 dari $\ln L_L(\gamma,\beta)$, jika hasilnya negatif, maka maksimum *likelihood* untuk γ dan β didapat dengan menyelesaikan persamaan $\frac{\partial \ln L_L(\gamma,\beta)}{\partial\gamma}=0$ dan $\frac{\partial \ln L_L(\gamma,\beta)}{\partial\beta}=0$.

Dari persamaan *likelihood* sampel tersensor tipe II diperoleh fungsi *likelihood* untuk distribusi log-logistik

$$L_L(\gamma,\beta)=\frac{n!}{(n-r)!}\left[\left(\frac{\beta}{\gamma}\right)^r\prod_{i=1}^{r}\frac{\left(\frac{t_i}{\gamma}\right)^{\beta-1}}{\left[1+\left(\frac{t_i}{\gamma}\right)^{\beta}\right]^2}\right]\left[\frac{1}{1+\left(\frac{t_r}{\gamma}\right)^{\beta}}\right]^{n-r}$$

Kemudian fungsi *likelihood* dikalikan dengan logaritma natural *(ln)*, sehingga diperoleh fungsi *log-likelihood* dari distribusi log-logistik sebagai berikut

$$\ln L_L(\gamma,\beta) = \ln\frac{n!}{(n-r)!} + \ln\beta^r - \ln\gamma^r + \sum_{i=1}^{r}\ln\left(\frac{t_i}{\gamma}\right)^{\beta-1} + \sum_{i=1}^{r}\ln\left[1+\left(\frac{t_i}{\gamma}\right)^{\beta}\right]^{-2} + \ln\left[1+\left(\frac{t_r}{\gamma}\right)^{\beta}\right]^{r-n}$$

$$\ln L_L(\gamma,\beta) = \ln\frac{n!}{(n-r)!} + \ln\beta^r - \ln\gamma^r + \sum_{i=1}^{r}(\beta-1)\ln\left(\frac{t_i}{\gamma}\right) + \sum_{i=1}^{r}-2\ln\left[1+\left(\frac{t_i}{\gamma}\right)^{\beta}\right]$$

$$+(r-n)\ln\left[1+\left(\frac{t_r}{\gamma}\right)^{\beta}\right] \tag{3.3}$$

Perhitungan turunan $\ln L_L(\gamma,\beta)$ terhadap γ adalah sebagai berikut

$$\frac{\partial \ln L_L(\gamma,\beta)}{\partial\gamma}$$

$$= -\frac{r}{\gamma} + \frac{r(1-\beta)}{\gamma} + (-2)\sum_{i=1}^{r}\left[1+\left(\frac{t_i}{\gamma}\right)^{\beta}\right]^{-1}\left(\frac{-\beta}{\gamma}\right)\left(\frac{t_i}{\gamma}\right)^{\beta} - \frac{(r-n)}{1+\left(\frac{t_r}{\gamma}\right)^{\beta}}\left(\frac{\beta}{\gamma}\right)\left(\frac{t_r}{\gamma}\right)^{\beta}$$

$$= \frac{-r\beta}{\gamma} + \left(\frac{2\beta}{\gamma}\right)\sum_{i=1}^{r}\left[1+\left(\frac{t_i}{\gamma}\right)^{\beta}\right]^{-1}\left(\frac{t_i}{\gamma}\right)^{\beta} - \frac{(r-n)}{1+\left(\frac{t_r}{\gamma}\right)^{\beta}}\left(\frac{\beta}{\gamma}\right)\left(\frac{t_r}{\gamma}\right)^{\beta} \tag{3.4}$$

Perhitungan turunan ke-2 dari $\ln L_L(\gamma,\beta)$ terhadap γ adalah sebagai berikut

$$\frac{\partial^2 \ln L_L(\gamma,\beta)}{\partial\gamma^2} = \frac{r\beta}{\gamma^2} - 2\left(\frac{\beta}{\gamma}\right)\left[\sum_{i=1}^{r}\frac{\left[\left(\frac{\beta+1}{\gamma}\right)\left(\frac{t_i}{\gamma}\right)^{\beta}\left(1+\left(\frac{t_i}{\gamma}\right)^{\beta}\right)-\left(\frac{\beta}{\gamma}\right)\left(\frac{t_i}{\gamma}\right)^{2\beta}\right]}{\left[1+\left(\frac{t_i}{\gamma}\right)^{\beta}\right]^2}\right]$$

$$-\frac{(r-n)}{\left[1+\left(\frac{t_r}{\gamma}\right)^{\beta}\right]^2}\left[\left(\frac{-\beta-1}{\gamma}\right)\left(\frac{\beta}{\gamma}\right)\left(\frac{t_r}{\gamma}\right)^{\beta}\left(1+\left(\frac{t_r}{\gamma}\right)^{\beta}\right)+\left(\frac{\beta}{\gamma}\right)^2\left(\frac{t_r}{\gamma}\right)^{2\beta}\right] \quad (3.5)$$

Perhitungan turunan ke-2 dari $\ln L_L(\gamma,\beta)$ terhadap γ mendapatkan hasil yang negatif, sehingga *maximum likelihood estimator* $\hat{\gamma}$ diperoleh dengan menyelesaikan $\frac{\partial \ln L_L(\gamma,\beta)}{\partial\gamma} = 0$

$$-\frac{r\hat{\beta}}{\hat{\gamma}} + \left(\frac{2\hat{\beta}}{\hat{\gamma}}\right)\sum_{i=1}^{r}\left[1+\left(\frac{t_i}{\hat{\gamma}}\right)^{\hat{\beta}}\right]^{-1}\left(\frac{t_i}{\hat{\gamma}}\right)^{\hat{\beta}} - \frac{(r-n)}{1+\left(\frac{t_r}{\hat{\gamma}}\right)^{\hat{\beta}}}\left(\frac{\hat{\beta}}{\hat{\gamma}}\right)\left(\frac{t_r}{\hat{\gamma}}\right)^{\hat{\beta}} = 0$$

$$(3.6)$$

Perhitungan turunan $\ln L_L(\gamma,\beta)$ terhadap β adalah sebagai berikut

$$\frac{\partial \ln L_L(\gamma,\beta)}{\partial\beta} = \frac{r}{\beta} + \sum_{i=1}^{r}\ln\frac{t_i}{\gamma} - 2\sum_{i=1}^{r}\left[1+\left(\frac{t_i}{\gamma}\right)^{\beta}\right]^{-1}\left(\frac{t_i}{\gamma}\right)^{\beta}\ln\left(\frac{t_i}{\gamma}\right) + \frac{(r-n)}{1+\left(\frac{t_r}{\gamma}\right)^{\beta}}\left(\frac{t_r}{\gamma}\right)^{\beta}\ln\left(\frac{t_r}{\gamma}\right)$$

$$(3.7)$$

Perhitungan turunan ke-2 dari $\ln L_L(\gamma,\beta)$ terhadap β adalah sebagai berikut

$$\frac{\partial^2 \ln L_L(\gamma,\beta)}{\partial \beta^2} = -\frac{r}{\beta^2} - 2\sum_{i=1}^{r} \frac{\left(\frac{t_i}{\gamma}\right)^{\beta}\left(\ln\left(\frac{t_i}{\gamma}\right)\right)^2}{\left[1+\left(\frac{t_i}{\gamma}\right)^{\beta}\right]^2} + \frac{(r-n)}{\left[1+\left(\frac{t_r}{\gamma}\right)^{\beta}\right]^2}\left(\frac{t_r}{\gamma}\right)^{\beta}\left(\ln\left(\frac{t_r}{\gamma}\right)\right)^2 \tag{3.8}$$

Perhitungan turunan ke-2 dari $\ln L_L(\gamma,\beta)$ terhadap β mendapatkan hasil yang negatif, sehingga *maximum likelihood estimator* $\hat{\beta}$ diperoleh dengan menyelesaikan $\frac{\partial \ln L_L(\gamma,\beta)}{\partial \beta} = 0$.

$$\frac{r}{\hat{\beta}} + \sum_{i=1}^{r} \ln\frac{t_i}{\hat{\gamma}} - 2\sum_{i=1}^{r}\left[1+\left(\frac{t_i}{\hat{\gamma}}\right)^{\hat{\beta}}\right]^{-1}\left(\frac{t_i}{\hat{\gamma}}\right)^{\hat{\beta}} \ln\left(\frac{t_i}{\hat{\gamma}}\right) + \frac{(r-n)}{1+\left(\frac{t_r}{\hat{\gamma}}\right)^{\hat{\beta}}}\left(\frac{t_r}{\hat{\gamma}}\right)^{\hat{\beta}} \ln\left(\frac{t_r}{\hat{\gamma}}\right) = 0$$

Jadi dari perhitungan yang telah dilakukan, maka *maximum likelihood estimator* $\hat{\gamma}$ dan $\hat{\beta}$ diperoleh dengan menyelesaikan persamaan

$$-\frac{r\hat{\beta}}{\hat{\gamma}} + \left(\frac{2\hat{\beta}}{\hat{\gamma}}\right)\sum_{i=1}^{r}\left[1+\left(\frac{t_i}{\hat{\gamma}}\right)^{\hat{\beta}}\right]^{-1}\left(\frac{t_i}{\hat{\gamma}}\right)^{\hat{\beta}} - \frac{(r-n)}{1+\left(\frac{t_r}{\hat{\gamma}}\right)^{\hat{\beta}}}\left(\frac{\hat{\beta}}{\hat{\gamma}}\right)\left(\frac{t_r}{\hat{\gamma}}\right)^{\hat{\beta}} = 0 \tag{3.9}$$

dan

$$\frac{r}{\hat{\beta}} + \sum_{i=1}^{r} \ln\frac{t_i}{\hat{\gamma}} - 2\sum_{i=1}^{r}\left[1+\left(\frac{t_i}{\hat{\gamma}}\right)^{\hat{\beta}}\right]^{-1}\left(\frac{t_i}{\hat{\gamma}}\right)^{\hat{\beta}} \ln\left(\frac{t_i}{\hat{\gamma}}\right) + \frac{(r-n)}{1+\left(\frac{t_r}{\hat{\gamma}}\right)^{\hat{\beta}}}\left(\frac{t_r}{\hat{\gamma}}\right)^{\hat{\beta}} \ln\left(\frac{t_r}{\hat{\gamma}}\right) = 0 \tag{3.10}$$

Kedua persamaan tersebut sulit diselesaikan secara manual karena memiliki bentuk yang kompleks, sehingga diperlukan bantuan dengan menggunakan suatu program atau *software* tertentu yang dapat digunakan untuk analisis *survival* dengan distribusi log-logistik.

D. Interval Konfidensi

Pada analisis *survival* ini, setelah didapatkan nilai dari γ dan β, selanjutnya akan dihitung interval konfidensi untuk γ dan β. Langkah pertama adalah menentukan matrik informasi $\left(I(\gamma,\beta)\right)$ dari data yaitu

$$I_{ij}(\theta) = E\left(\frac{-\partial^2 \log L(\theta)}{\partial\theta_i \partial\theta_j}\right)$$

$$I(\gamma,\beta) = \begin{bmatrix} -\dfrac{\partial^2 \ln L}{\partial\gamma^2} & -\dfrac{\partial^2 \ln L}{\partial\gamma\partial\beta} \\ -\dfrac{\partial^2 \ln L}{\partial\beta\partial\gamma} & -\dfrac{\partial^2 \ln L}{\partial\beta^2} \end{bmatrix}$$

Dari perhitungan sebelumnya telah didapatkan persamaan $\dfrac{\partial^2 \ln L}{\partial\gamma^2}$ dan $\dfrac{\partial^2 \ln L}{\partial\beta^2}$ pada persamaan (3.5) dan (3.8). selanjutnya akan dihitung persamaan $\dfrac{\partial^2 \ln L}{\partial\gamma\partial\beta}$ dan $\dfrac{\partial^2 \ln L}{\partial\beta\partial\gamma}$. Kedua persamaan tersebut memiliki hasil yang sama, sehingga

$$\frac{\partial^2 \ln L}{\partial\gamma\partial\beta} = \frac{\partial^2 \ln L}{\partial\beta\partial\gamma}$$

$$= -\frac{r}{\gamma} + \frac{2}{\gamma}\left[\sum_{i=1}^{r}\left[\left(\frac{t_i}{\gamma}\right)^{\beta}\left[1+\beta\ln\left(\frac{t_i}{\gamma}\right)\right]\left[1+\left(\frac{t_i}{\gamma}\right)^{\beta}\right]-\beta\left(\frac{t_i}{\gamma}\right)^{2\beta}\ln\left(\frac{t_i}{\gamma}\right)\right]\left[1+\left(\frac{t_i}{\gamma}\right)^{\beta}\right]^{-2}\right]$$

$$-\frac{(r-n)}{\left[1+\left(\frac{t_r}{\gamma}\right)^{\beta}\right]^{2}}\left[\left(\frac{1}{\gamma}\right)\left(\frac{t_r}{\gamma}\right)^{\beta}\left[1+\beta\ln\left(\frac{t_r}{\gamma}\right)\right]\left[1+\left(\frac{t_r}{\gamma}\right)^{\beta}\right]-\left(\frac{\beta}{\gamma}\right)\left(\frac{t_r}{\gamma}\right)^{2\beta}\ln\left(\frac{t_r}{\gamma}\right)\right] \tag{3.11}$$

Matriks kovarian V merupakan invers dari matriks Informasi yaitu:

$$\hat{V} = I(\gamma,\beta)^{-1}$$

$$\hat{V} = \begin{bmatrix} \widehat{V_{11}} & \widehat{V_{12}} \\ \widehat{V_{21}} & \widehat{V_{22}} \end{bmatrix}$$

dengan *standard error* untuk $\hat{\gamma}$ adalah $se(\hat{\gamma}) = \hat{V}_{11}^{1/2}$, dan standard error untuk $\hat{\beta}$ adalah $se(\hat{\beta}) = \hat{V}_{22}^{1/2}$.

Interval konfidensi untuk γ dan β dapat diperoleh dari pendekatan nilai pivot

$$Z_1 = \frac{\hat{\gamma}-\gamma}{se(\hat{\gamma})} \text{ dan } Z_2 = \frac{\hat{\beta}-\beta}{se(\hat{\beta})}$$

dengan keduanya mendekati distribusi normal *N(0,1)* untuk sampel besar. Sehingga dengan taraf signifikansi $1-\alpha$ didapatkan interval konfidensi untuk γ adalah

$$P(-z_{\alpha/2} \le Z_1 \le z_{\alpha/2}) = 1-\alpha$$

$$P(-z_{\alpha/2} \le \frac{\hat{\gamma}-\gamma}{se(\hat{\gamma})} \le z_{\alpha/2}) = 1-\alpha$$

$$\hat{\gamma} - z_{\alpha/2} se(\hat{\gamma}) \le \gamma \le \hat{\gamma} + z_{\alpha/2} se(\hat{\gamma})$$

dan interval konfidensi $1-\alpha$ untuk β adalah

$$P(-z_{\alpha/2} \le Z_2 \le z_{\alpha/2}) = 1-\alpha$$

$$P(-z_{\alpha/2} \le \frac{\hat{\beta}-\beta}{se(\hat{\beta})} \le z_{\alpha/2}) = 1-\alpha$$

$$\hat{\beta} - z_{\alpha/2} se(\hat{\beta}) \le \beta \le \hat{\beta} + z_{\alpha/2} se(\hat{\beta}).$$

E. Contoh Penerapan

Berikut ini adalah data 30 umur penyakit hingga pasien meninggal dari 50 pasien yang menderita penyakit kanker paru-paru. Data berasal dari hasil metode simulasi pembangkitan data dengan bantuan *software Minitab* 14 yang berdistribusi log-logistik.

Tabel 3.1. Data umur penyakit pasien (bulan)

20,835	36,917	45,787	61,168	71,817
23,364	37,794	46,238	64,449	72,157
27,959	40,327	49,080	64,562	72,896
30,830	41,869	53,179	65,448	72,992
31,395	42,985	56,004	67,540	73,202
33,600	43,959	59,128	69,055	74,316

Dari 50 pengamatan yang ada, hanya diambil 30 hasil pengamatan pertama. Banyaknya pengamatan yang diambil telah ditentukan sebelum penelitian dilakukan. Pada data ini hanya diambil 30 pengamatan, sehingga terdapat 20 pengamatan yang tersensor. Akan dicari nilai dari *maximum likelihood estimator* untuk γ dan β, untuk menghitung peluang hidup seorang pasien yang menderita penyakit kanker paru-paru selama 50 dan 90 bulan.

Untuk mempermudah perhitungan dalam mencari nilai *maximum likelihood estimator*, dapat menggunakan *software Minitab*. Dalam skripsi ini menggunakan *software Minitab* 14. Dalam *software* ini, fungsi densitas peluang yang digunakan adalah

$$f(y) = \frac{\exp\left(\dfrac{\ln y - \mu}{\sigma}\right)}{\sigma\left[1 + \exp\left(\dfrac{\ln y - \mu}{\sigma}\right)\right]^2} \tag{3.12}$$

dengan $\mu =$ *location parameter*

$\sigma =$ *scale parameter.*

Bentuk fungsi densitas pada persamaan (3.12) merupakan hasil transformasi dari fungsi densitas

$$f(t;\gamma,\beta) = \frac{\left(\dfrac{\beta}{\gamma}\right)\left(\dfrac{t}{\gamma}\right)^{\beta-1}}{\left[1+\left(\dfrac{t}{\gamma}\right)^{\beta}\right]^2}, \text{ dimana } \gamma = \exp(\mu) \text{ dan } \beta = \frac{1}{\sigma}.$$

Dari hasil output *software Minitab* 14 (lampiran 1), diperoleh nilai *location parameter* dari data adalah 4,37719 dan nilai dari *scale parameter* adalah 0,465591, sehingga didapatkan:

$\hat{\gamma}$ = *exp(location parameter)*

$= e^{4,37719}$

$= 79,61400344$

$$\widehat{\beta} = 1/\ scale\ parameter$$

$$= \frac{1}{0,465591}$$

$$= 2,14780784$$

Setelah dilakukan pengecekan dengan *software maple 11* (lampiran 2), nilai dari $\hat{\gamma}$ dan $\widehat{\beta}$ tersebut memenuhi persamaan (3.9) dan (3.10), karena menghasilkan nilai yang mendekati nol. Jadi nilai untuk $\hat{\gamma}$ adalah $79,61400344$ dan nilai untuk $\widehat{\beta}$ adalah $2,14780784$.

Selanjutnya akan ditentukan interval konfidensi untuk γ dan β . Dari perhitungan sebelumnya telah didapatkan persamaan $\frac{\partial^2 \ln L}{\partial \gamma^2}$, $\frac{\partial^2 \ln L}{\partial \beta^2}$ dan $\frac{\partial^2 \ln L}{\partial \gamma \partial \beta} = \frac{\partial^2 \ln L}{\partial \beta \partial \gamma}$ pada persamaan (3.5), (3.8), dan (3.11). Dengan bantuan program *maple 11* pada lampiran 3, nilai dari $\frac{\partial^2 \ln L}{\partial \gamma^2}$ dimana $\hat{\gamma} = 79,61400344$ dan $\widehat{\beta} = 2,1478784$ adalah $-0,01\,047404785$. Nilai dari $\frac{\partial^2 \ln L}{\partial \beta^2}$ pada lampiran 4 adalah $-\ 9,068024511$, sedangkan nilai dari $\frac{\partial^2 \ln L}{\partial \gamma \partial \beta} = \frac{\partial^2 \ln L}{\partial \beta \partial \gamma}$ pada lampiran 5 adalah -0,1680850939. Oleh karena itu didapatkan matrik informasi

$$I(\gamma, \beta) = \begin{bmatrix} -\frac{\partial^2 \ln L}{\partial \gamma^2} & -\frac{\partial^2 \ln L}{\partial \gamma \partial \beta} \\ -\frac{\partial^2 \ln L}{\partial \beta \partial \gamma} & -\frac{\partial^2 \ln L}{\partial \beta^2} \end{bmatrix}$$

$$I(\gamma,\beta)=\begin{bmatrix}0,0104740478 & 0,1680850939\\ 0,1680850939 & 9,0680245110\end{bmatrix}$$

dan matrik kovarian

$$\hat{V}=I(\gamma,\beta)^{-1}$$

$$\hat{V}=\begin{bmatrix}0,0104740478 & 0,1680850939\\ 0,1680850939 & 9,0680245110\end{bmatrix}^{-1}$$

$$\hat{V}=\begin{bmatrix}\hat{V}_{11} & \hat{V}_{12}\\ \hat{V}_{21} & \hat{V}_{22}\end{bmatrix}$$

$$\hat{V}=\begin{bmatrix}135,8987585 & -2,519022259\\ -2,519022259 & 0,156970252\end{bmatrix},$$

sehingga $se(\hat{\gamma})=\hat{V}_{11}^{1/2}=11,65756229$ dan $se(\hat{\beta})=\hat{V}_{22}^{1/2}=0,396194714$.

Interval konfidensi untuk γ didapatkan dari nilai pivot $Z_1=\dfrac{\hat{\gamma}-\gamma}{se(\hat{\gamma})}$ yang berdistribusi normal. Dengan taraf signifikansi 0,05, interval konfidensi untuk γ yaitu $\hat{\gamma}\pm z_{0,025}se(\hat{\gamma})$. Dari tabel z pada lampiran 6 didapatkan nilai dari $z_{0,025}$ adalah 1,96, dengan demikian

$$\hat{\gamma}-z_{\alpha/2}se(\hat{\gamma})\le\gamma\le\hat{\gamma}+z_{\alpha/2}se(\hat{\gamma})$$

$$\hat{\gamma}-1,96\,se(\hat{\gamma})\le\gamma\le\hat{\gamma}+1,96\,se(\hat{\gamma})$$

$$79,61400344-1,96(11,65756229)\le\gamma\le 79,61400344+1,96(11,65756229)$$

$$56,76518135\le\gamma\le 102,4628255.$$

Jadi batas konfidensi bawah untuk γ adalah 56,76518135 dan batas konfidensi atasnya adalah 102,4628255.

Interval konfidensi untuk β didapatkan dari nilai pivot $Z_2 = \dfrac{\widehat{\beta} - \beta}{se(\widehat{\beta})}$ yang berdistribusi normal. Dengan taraf signifikansi 0,05, interval konfidensi untuk β yaitu $\widehat{\beta} \pm z_{\alpha/2} se(\widehat{\beta})$. Dari tabel z pada lampiran 6 didapatkan nilai dari $z_{0,025}$ adalah 1,96, dengan demikian

$$\widehat{\beta} - z_{\alpha/2} se(\widehat{\beta}) \le \beta \le \widehat{\beta} + z_{\alpha/2} se(\widehat{\beta})$$

$$\widehat{\beta} - 1,96\, se(\widehat{\beta}) \le \beta \le \widehat{\beta} + 1,96\, se(\widehat{\beta})$$

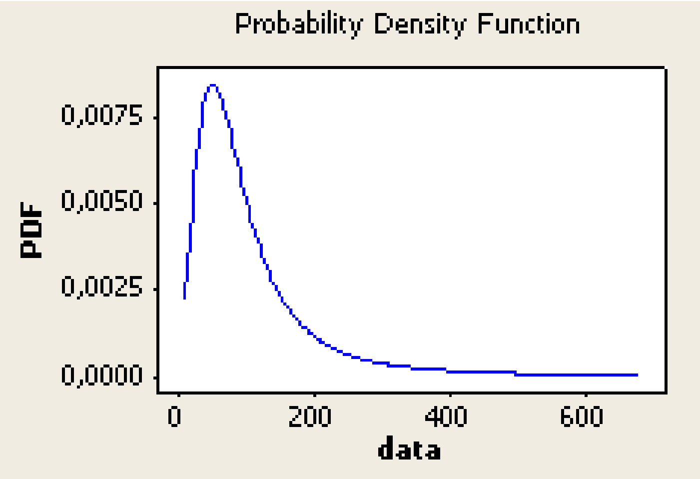
$$2,1478784 - 1,96(0,396194714) \le \beta \le 2,1478784 + 1,96(0,396194714)$$

$$1,371336761 \le \beta \le 2,924420039$$

Jadi batas konfidensi bawah untuk β adalah 1.371336761 dan batas konfidensi atasnya adalah 2.924420039.

Berikut ini adalah bentuk fungsi densitas peluang dari data, sedangkan bentuk kurvanya ditunjukkan pada gambar 3.2:

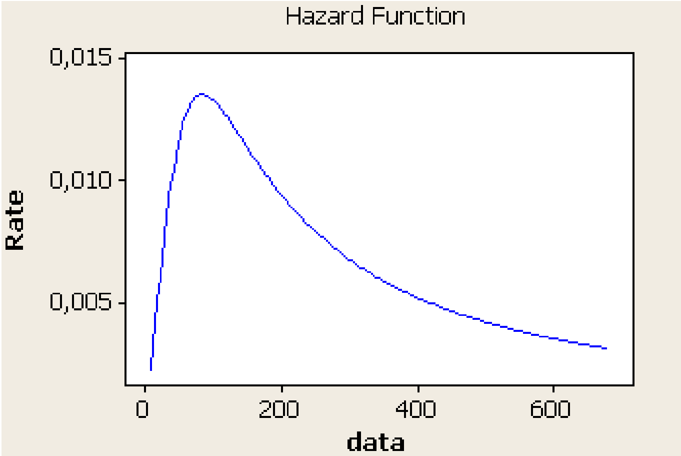
$$f(t_i) = \frac{(0,02697776455)\left(\dfrac{t_i}{79,61400344}\right)^{1,14780784}}{\left[1 + \left(\dfrac{t_i}{79,61400344}\right)^{2,14780784}\right]^2} \qquad (3.13)$$



Gambar 3.2. Kurva fungsi densitas peluang dari data umur penyakit pasien

Bentuk fungsi *hazard* ditunjukkan pada persamaan 3.14 dan kurva fungsi *hazard* dari data pada gambar 3.3.

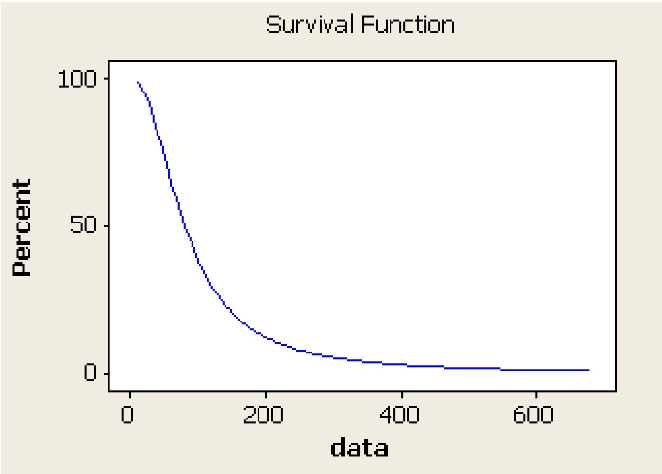
$$h(t_i) = \frac{(0,02697776455)\left(\frac{t_i}{79,61400344}\right)^{1,14780784}}{\left[1+\left(\frac{t_i}{79,61400344}\right)^{2,14780784}\right]} \tag{3.14}$$



Gambar 3.3 Kurva fungsi *hazard* dari data umur penyakit pasien

Bentuk fungsi *survivor* ditunjukkan pada persamaan 3.15 dan kurva fungsi *survivor* dari data pada gambar 3.4.

$$S(t_i) = \frac{1}{1 + \left(\dfrac{t_i}{79,61400344} \right)^{2,14780784}} \qquad (3.15)$$



Gambar 3.4 Kurva fungsi *survivor* dari data umur penyakit pasien

Jika fungsi *survivor* telah didapatkan, maka dapat dihitung peluang seorang pasien untuk hidup jika menderita penyakit kanker paru-paru selama 50 dan 90 bulan.

$$S(50) = \frac{1}{1 + \left(\dfrac{50}{79,61400344} \right)^{2,14780784}}$$

$$S(50) = 0,73087949$$

Jadi peluang seorang pasien untuk hidup jika menderita kanker paru-paru selama 50 bulan adalah 0,73. Sedangkan peluang pasien untuk hidup jika menderita kanker paru-paru selama 90 bulan adalah

$$S(90) = \frac{1}{1 + \left(\frac{90}{79,61400344} \right)^{2,14780784}}$$

$$S(90) = 0,4345370684.$$

Jadi peluang seorang pasien untuk hidup jika menderita kanker paru-paru selama 90 bulan adalah 0,43.

BAB IV

PENUTUP

A. KESIMPULAN

Dari pembahasan, diperoleh beberapa kesimpulan mengenai model *survival* dan inferensia statistik data tahan hidup tersensor tipe II yang berdistribusi log-logistik, yaitu sebagai berikut:

1. Model *survival* untuk data tersensor tipe II berdasarkan model distribusi log-logistik adalah

$$L_L(\gamma,\beta)=\frac{n!}{(n-r)!}\left[\left(\frac{\beta}{\gamma}\right)^r\prod_{i=1}^{r}\frac{\left(\frac{t_i}{\gamma}\right)^{\beta-1}}{\left[1+\left(\frac{t_i}{\gamma}\right)^{\beta}\right]^2}\right]\left[\frac{1}{1+\left(\frac{t_r}{\gamma}\right)^{\beta}}\right]^{n-r}.$$

   Model *survival* ini diperoleh dengan mencari fungsi likelihood dari data tersensor tipe II yang berdistribusi log-logistik. Fungsi likelihood untuk data tersensor tipe II tersebut adalah $f(t_1,t_2,...,t_r)=\dfrac{n!}{(n-r)!}\left[\prod_{i=1}^{r}f(t_i)\right][S(t_r)]^{n-r}$

2. Inferensia statistik data *survival* tersensor tipe II berdasarkan distribusi log-logistik adalah sebagai berikut:

   a. Menentukan nilai *maximum likelihood estimator* untuk γ dan β yaitu $\hat{\gamma}$ dan $\hat{\beta}$. Dari fungsi likelihood untuk data tersensor tipe II, dapat diperoleh *maximum likelihood estimator* untuk γ dan β yaitu $\hat{\gamma}$ dan $\hat{\beta}$ dengan menurunkan fungsi log-likelihoodnya terhadap γ dan terhadap β dan menyelesaikan kedua persamaan tersebut, yaitu

$$\frac{1-\hat{\beta}-r}{\hat{\gamma}}+\left(\frac{2\hat{\beta}}{\hat{\gamma}}\right)\sum_{i=1}^{r}\left[1+\left(\frac{t_i}{\hat{\gamma}}\right)^{\hat{\beta}}\right]^{-1}\left(\frac{t_i}{\hat{\gamma}}\right)^{\hat{\beta}}-\frac{(r-n)}{1+\left(\frac{t_r}{\hat{\gamma}}\right)^{\hat{\beta}}}\left(\frac{\hat{\beta}}{\hat{\gamma}}\right)\left(\frac{t_r}{\hat{\gamma}}\right)^{\hat{\beta}}=0$$

dan

$$\frac{r}{\hat{\beta}}+\sum_{i=1}^{r}\ln\frac{t_i}{\hat{\gamma}}-2\sum_{i=1}^{r}\left[1+\left(\frac{t_i}{\hat{\gamma}}\right)^{\hat{\beta}}\right]^{-1}\left(\frac{t_i}{\hat{\gamma}}\right)^{\hat{\beta}}\ln\left(\frac{t_i}{\hat{\gamma}}\right)+\frac{(r-n)}{1+\left(\frac{t_r}{\hat{\gamma}}\right)^{\hat{\beta}}}\left(\frac{t_r}{\hat{\gamma}}\right)^{\hat{\beta}}\ln\left(\frac{t_r}{\hat{\gamma}}\right)=0$$

Kedua persamaan tersebut sulit diselesaikan secara manual karena memiliki bentuk yang kompleks, sehingga dalam skripsi ini memerlukan bantuan *software Minitab14* untuk mencari nilai *maximum likelihood estimatornya*.

b. Menentukan interval konfidensi untuk γ dan β.

Untuk menentukan interval konfidensi data tersensor , digunakan pendekatan nilai pivot $Z_1=\frac{\hat{\gamma}-\gamma}{se(\hat{\gamma})}$ dan $Z_1=\frac{\hat{\beta}-\beta}{se(\hat{\beta})}$ dengan keduanya mendekati distribusi normal N(0,1). Oleh karena itu interval konfidensi untuk γ adalah $\hat{\gamma}-z_{\alpha/2}se(\hat{\gamma})\le\gamma\le\hat{\gamma}+z_{\alpha/2}se(\hat{\gamma})$ dan interval konfidensi untuk β adalah $\hat{\beta}-z_{\alpha/2}se(\hat{\beta})\le\beta\le\hat{\beta}+z_{\alpha/2}se(\hat{\beta})$.

3. Hasil pengolahan dari data umur pasien yang menderita penyakit kanker paru-paru dengan simulasi pembangkitan data berdistribusi log-logistik adalah nilai estimasi untuk γ yaitu 79,61400344 dan estimasi untuk β yaitu 2,14780784. Sedangkan interval konfidensi untuk γ adalah $56,76518135\le\gamma\le102,4628255$ dan interval konfidensi untuk β adalah

$1,371336761 \le \beta \le 2,924420039$. Sehingga didapatkan peluang hidup untuk pasien yang menderita kanker paru-paru selama 50 bulan adalah 0,73, sedangkan peluang hidup pasien yang menderita selama 90 bulan adalah 0,43.

B. SARAN

Skripsi ini membahas tentang model *survival* dengan menentukan *maximum likelihood estimator* untuk γ dan β yang merupakan parameter-parameter dari distribusi log-logistik. Dalam penulisan ini hanya membahas model *survival* untuk data tersensor tipe II. Oleh karena itu disarankan adanya penelitian lebih lanjut mengenai model *survival* dengan menggunakan distribusi log-logistik untuk data tersensor tipe I dan juga untuk distribusi-distribusi lain pada data berkelompok.

DAFTAR PUSTAKA

Abadyo dan Hendro Permadi. 2005. *Metode Statistika Praktis*. Malang: UM Press.

Bain, L.J and Engelhardt. 1992. *Introduction to Probability and Mathematical Statistics*. 2nd ed. California: Duxbury Press.

Collett, David. 2004. *Modelling Survival Data in Medical Research*. 2nd ed. London: Chapman and Hall.

Dixit, Asha. 2008. *Exact Comparison of Hazard Rate Functions of Log-logistic Survival Distribution* [Tesis]. Alabama: Auburn University.

Lawless, J.F. 1982. *Statistical Model and Methods for Lifetime Data*. New York: John Wiley and Sons, Inc.

Lawless, J.F. 2003. *Statistical Model and Methods for Lifetime Data.* 2nd ed. New Jersey: John Wiley and Sons Inc.

Machin, David, Yin Bun C and Mahesh Parmar. 2006. *Survival Analysis A Practical Approach.* 2nd ed. Chicester: John Wiley and Sons Ltd.

Rao, G.S, Kantam and K.Rosaih. 2009. "Reliability Estimation in Log-logistic Distribution from Cencored Samples", *Prob.Stat*.,02,52-67.

Walpole, Ronald E. 1993. *Pengantar Statistika Edisi ke-3*. Jakarta: Gramedia Pustaka Utama.

LAMPIRAN

Lampiran 1

Output Hasil Analisis Survival Menggunakan *Minitab 14*

Distribution Analysis: data

```
Variable: data

Censoring Information  Count
Uncensored value          30
Right censored value      20

Type 2 (Failure) Censored at 31

Estimation Method: Maximum Likelihood

Distribution:   Loglogistic


Parameter Estimates

                      Standard    95,0% Normal CI
Parameter  Estimate      Error     Lower     Upper
Location    4,37719   0,122486   4,13712   4,61726
Scale      0,465591  0,0725429  0,343068  0,631871

Log-Likelihood = -168,968

Goodness-of-Fit
Anderson-Darling (adjusted) = 128,746
```

Lampiran 2

Output Hasil Perhitungan Menggunakan *Maple 11*

> $t := array(1..30, [20.835, 23.364, 27.959, 30.830, 31.395, 33.6, 36.917, 37.794, 40.327, 41.869, 42.985, 43.959, 45.787, 46.238, 49.080, 53.179, 56.004, 59.128, 61.168, 64.449, 64.562, 65.448, 67.540, 69.055, 71.817, 72.157, 72.896, 72.992, 73.202, 74.316])$

$t := [20.835, 23.364, 27.959, 30.830, 31.395, 33.6, 36.917, 37.794, 40.327, 41.869, 42.985, 43.959, 45.787, 46.238, 49.080, 53.179, 56.004, 59.128, 61.168, 64.449, 64.562, 65.448, 67.540, 69.055, 71.817, 72.157, 72.896, 72.992, 73.202, 74.316]$

=

$$> f1 := \left(-\frac{30\cdot\frac{1}{0.465591}}{79.61400344}\right) + \left(2\cdot\frac{1}{0.465591}\cdot\frac{1}{79.61400344}\right)\cdot\left(sum\left(\frac{\left(\frac{(t[i])}{79.61400344}\right)^{\frac{1}{0.465591}}}{1+\left(\frac{(t[i])}{79.61400344}\right)^{\frac{1}{0.465591}}}, i = 1..30\right)\right) + 20\cdot\left(\frac{1}{0.465591}\cdot\frac{1}{79.61400344}\right)\cdot\left(\frac{\left(\frac{74.316}{79.61400344}\right)^{\frac{1}{0.465591}}}{1+\left(\frac{74.316}{79.61400344}\right)^{\frac{1}{0.465591}}}\right)$$

$$f1 := -0.0999940370$$

=

$$> f2 := 30\cdot 0.465591 + sum\left(ln\left(\frac{t[i]}{79.61400344}\right), i = 1..30\right) - 2\cdot sum\left(\frac{\left(\frac{(t[i])}{79.61400344}\right)^{\frac{1}{0.465591}}\cdot\ln\left(\frac{(t[i])}{79.61400344}\right)}{1+\left(\frac{(t[i])}{79.61400344}\right)^{\frac{1}{0.465591}}}, i = 1..30\right) - 20\cdot\frac{\left(\frac{74.316}{79.61400344}\right)^{\frac{1}{0.465591}}\cdot\ln\left(\frac{74.316}{79.61400344}\right)}{1+\left(\frac{74.316}{79.61400344}\right)^{\frac{1}{0.465591}}}$$

$$f2 := 0.7579785301$$

=

Lampiran 3

Output Hasil Perhitungan Menggunakan *Maple 11*

> $t := array(1..30, [20.835, 23.364, 27.959, 30.830, 31.395, 33.6, 36.917, 37.794, 40.327, 41.869, 42.985, 43.959, 45.787, 46.238, 49.080, 53.179, 56.004, 59.128, 61.168, 64.449, 64.562, 65.448, 67.540, 69.055, 71.817, 72.157, 72.896, 72.992, 73.202, 74.316])$

$$t := [20.835, 23.364, 27.959, 30.830, 31.395, 33.6, 36.917, 37.794, 40.327, 41.869, 42.985, 43.959, 45.787, 46.238, 49.080, 53.179, 56.004, 59.128, 61.168, 64.449, 64.562, 65.448, 67.540, 69.055, 71.817, 72.157, 72.896, 72.992, 73.202, 74.316] \qquad (1)$$

> $$\gamma 2 := \frac{\frac{1}{0.465591}\cdot 30}{79.61400344^2} - \frac{2\cdot\frac{1}{0.465591}}{79.61400344}\cdot sum\left(\frac{1}{\left(1+\left(\frac{(t[i])}{79.61400344}\right)^{\frac{1}{0.465591}}\right)^2}\left(\left(\frac{\frac{1}{0.465591}+1}{79.61400344}\right)\cdot\left(\frac{(t[i])}{79.61400344}\right)^{\frac{1}{0.465591}}\cdot\left(1+\left(\frac{(t[i])}{79.61400344}\right)^{\frac{1}{0.465591}}\right)-\left(\frac{1}{0.465591}\cdot\frac{1}{79.61400344}\right)\cdot\left(\frac{(t[i])}{79.61400344}\right)^{\frac{2}{0.465591}}\right), i=1..30\right) + \frac{20}{\left(1+\left(\frac{74.316}{79.61400344}\right)^{\frac{1}{0.465591}}\right)^2}\cdot\left(\left(\frac{-\frac{1}{0.465591}-1}{79.61400344}\right)\cdot\left(\frac{1}{0.465591}\cdot\frac{1}{79.61400344}\right)\cdot\left(\frac{74.316}{79.61400344}\right)^{\frac{1}{0.465591}}\cdot\left(1+\left(\frac{74.316}{79.61400344}\right)^{\frac{1}{0.465591}}\right)+\left(\frac{1}{0.465591}\cdot\frac{1}{79.61400344}\right)^2\cdot\left(\frac{74.316}{79.61400344}\right)^{\frac{2}{0.465591}}\right)$$

$$\gamma 2 := -0.01047404785 \qquad (2)$$

Lampiran 4

Output Hasil Perhitungan Menggunakan *Maple 11*

> $t := array(1..30, [20.835, 23.364, 27.959, 30.830, 31.395, 33.6, 36.917, 37.794, 40.327, 41.869, 42.985, 43.959, 45.787, 46.238, 49.080, 53.179, 56.004, 59.128, 61.168, 64.449, 64.562, 65.448, 67.540, 69.055, 71.817, 72.157, 72.896, 72.992, 73.202, 74.316])$

$$t := [20.835, 23.364, 27.959, 30.830, 31.395, 33.6, 36.917, 37.794, 40.327, 41.869, 42.985, 43.959, 45.787, 46.238, 49.080, 53.179, 56.004, 59.128, 61.168, 64.449, 64.562, 65.448, 67.540, 69.055, 71.817, 72.157, 72.896, 72.992, 73.202, 74.316] \quad (1)$$

> $$\beta 2 := -\frac{30}{\left(\frac{1}{0.465591}\right)^2} - 2 \cdot sum\left(\frac{\left(\frac{(t[i])}{79.61400344}\right)^{\frac{1}{0.465591}} \cdot \left(\ln\left(\frac{(t[i])}{79.61400344}\right)\right)^2}{\left(1+\left(\frac{(t[i])}{79.61400344}\right)^{\frac{1}{0.465591}}\right)^2}, i = 1..30\right) - \frac{20}{\left(1+\left(\frac{74.316}{79.61400344}\right)^{\frac{1}{0.465591}}\right)^2} \cdot \left(\frac{74.316}{79.61400344}\right)^{\frac{1}{0.465591}} \cdot \left(\ln\left(\frac{74.316}{79.61400344}\right)\right)^2$$

$$\beta 2 := -9.068024511 \quad (2)$$

>

Lampiran 5

Output Hasil Perhitungan Menggunakan *Maple 11*

> $t := array(1..30, [20.835, 23.364, 27.959, 30.830, 31.395, 33.6, 36.917, 37.794, 40.327, 41.869, 42.985, 43.959, 45.787, 46.238, 49.080, 53.179, 56.004, 59.128, 61.168, 64.449, 64.562, 65.448, 67.540, 69.055, 71.817, 72.157, 72.896, 72.992, 73.202, 74.316])$

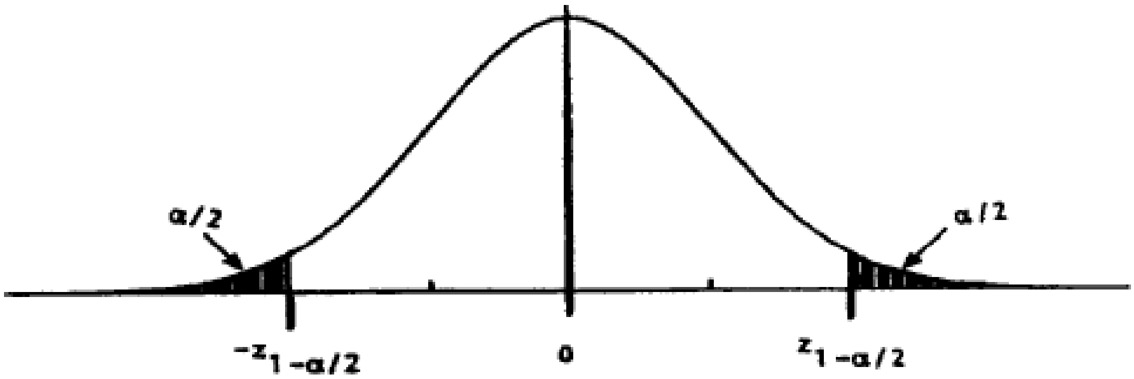
$$t := [20.835, 23.364, 27.959, 30.830, 31.395, 33.6, 36.917, 37.794, 40.327, 41.869, 42.985, 43.959, 45.787, 46.238, 49.080, 53.179, 56.004, 59.128, 61.168, 64.449, 64.562, 65.448, 67.540, 69.055, 71.817, 72.157, 72.896, 72.992, 73.202, 74.316] \tag{1}$$

> $$\gamma\beta := -\frac{30}{79.61400344} + (-2)\cdot sum\left(\left(\left(-\frac{1}{79.61400344}\right)\cdot\left(\frac{(t[i])}{79.61400344}\right)^{\frac{1}{0.465591}}\cdot\left(\left(\frac{1}{0.465591}\cdot\ln\left(\frac{(t[i])}{79.61400344}\right)+1\right)\cdot\left(1+\left(\frac{(t[i])}{79.61400344}\right)^{\frac{1}{0.465591}}\right)-\frac{1}{0.465591}\cdot\left(\frac{(t[i])}{79.61400344}\right)^{\frac{1}{0.465591}}\cdot\ln\left(\frac{(t[i])}{79.61400344}\right)\right)\right)\cdot\left(1+\left(\frac{(t[i])}{79.61400344}\right)^{\frac{1}{0.465591}}\right)^{-2}, i=1..30\right) + (-20)\cdot\left(-\frac{1}{79.61400344}\right)\cdot\left(\frac{74.316}{79.61400344}\right)^{\frac{1}{0.465591}}\cdot\left(\left(\frac{1}{0.465591}\cdot\ln\left(\frac{74.316}{79.61400344}\right)+1\right)\cdot\left(1+\left(\frac{74.316}{79.61400344}\right)^{\frac{1}{0.465591}}\right)-\frac{1}{0.465591}\cdot\left(\frac{74.316}{79.61400344}\right)^{\frac{1}{0.465591}}\cdot\ln\left(\frac{74.316}{79.61400344}\right)\right)\cdot\left(1+\left(\frac{74.316}{79.61400344}\right)^{\frac{1}{0.465591}}\right)^{-2}$$

$$\gamma\beta := -.1680850939 \tag{2}$$

Lampiran 6

Tabel Distribusi Normal



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.00	1.0000	0.9920	0.9840	0.9761	0.9681	0.9601	0.9522	0.9442	0.9362	0.9283
0.10	0.9203	0.9124	0.9045	0.8966	0.8887	0.8808	0.8729	0.8650	0.8572	0.8493
0.20	0.8415	0.8337	0.8259	0.8181	0.8103	0.8026	0.7949	0.7872	0.7795	0.7718
0.30	0.7642	0.7566	0.7490	0.7414	0.7339	0.7263	0.7188	0.7114	0.7039	0.6965
0.40	0.6892	0.6818	0.6745	0.6672	0.6599	0.6527	0.6455	0.6384	0.6312	0.6241
0.50	0.6171	0.6101	0.6031	0.5961	0.5892	0.5823	0.5755	0.5687	0.5619	0.5522
0.60	0.5485	0.5419	0.5353	0.5287	0.5222	0.5157	0.5093	0.5029	0.4965	0.4902
0.70	0.4839	0.4777	0.4715	0.4654	0.4593	0.4533	0.4473	0.4413	0.4354	0.4295
0.80	0.4237	0.4179	0.4122	0.4065	0.4009	0.3953	0.3898	0.3843	0.3789	0.3735
0.90	0.3681	0.3628	0.3576	0.3524	0.3472	0.3421	0.3371	0.3320	0.3271	0.3222
1.00	0.3173	0.3125	0.3077	0.3030	0.2983	0.2937	0.2891	0.2846	0.2801	0.2757

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
1.10	0.3173	0.3125	0.3077	0.3030	0.2983	0.2937	0.2891	0.2846	0.2801	0.2757
1.10	0.2713	0.2670	0.2627	0.2585	0.2543	0.2501	0.2460	0.2420	0.2380	0.2340
1.20	0.2301	0.2263	0.2225	0.2187	0.2150	0.2113	0.2077	0.2041	0.2005	0.1971
1.30	0.1936	0.1902	0.1868	0.1835	0.1802	0.1770	0.1738	0.1707	0.1676	0.1645
1.40	0.1615	0.1585	0.1556	0.1527	0.1499	0.1471	0.1443	0.1416	0.1389	0.1362
1.50	0.1336	0.1310	0.1285	0.1260	0.1236	0.1211	0.1188	0.1164	0.1141	0.1118
1.60	0.1096	0.1074	0.1052	0.1031	0.1010	0.0989	0.0969	0.0949	0.0930	0.0910
1.70	0.0891	0.0873	0.0854	0.0836	0.0819	0.0801	0.0784	0.0767	0.0751	0.0735
1.80	0.0719	0.0703	0.0688	0.0672	0.0658	0.0643	0.0629	0.0615	0.0601	0.0588
1.90	0.0574	0.0561	0.0549	0.0536	0.0524	0.0512	0.0500	0.0488	0.0477	0.0466
2.00	0.0455	0.0444	0.0434	0.0424	0.0414	0.0404	0.0394	0.0385	0.0375	0.0366

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
2.00	0.0455	0.0444	0.0434	0.0424	0.0414	0.0404	0.0394	0.0385	0.0375	0.0366
2.10	0.0357	0.0349	0.0340	0.0332	0.0324	0.0316	0.0308	0.0300	0.0293	0.0285
2.20	0.0278	0.0271	0.0264	0.0257	0.0251	0.0244	0.0238	0.0232	0.0226	0.0220
2.30	0.0214	0.0209	0.0203	0.0198	0.0193	0.0188	0.0183	0.0178	0.0173	0.0168
2.40	0.0164	0.0160	0.0155	0.0151	0.0147	0.0143	0.0139	0.0135	0.0131	0.0128
2.50	0.0124	0.0121	0.0117	0.0114	0.0111	0.0108	0.0105	0.0102	0.0099	0.0096
2.60	0.0093	0.0091	0.0088	0.0085	0.0083	0.0080	0.0078	0.0076	0.0074	0.0071
2.70	0.0069	0.0067	0.0065	0.0063	0.0061	0.0060	0.0058	0.0056	0.0054	0.0053
2.80	0.0051	0.0050	0.0048	0.0047	0.0045	0.0044	0.0042	0.0041	0.0040	0.0039
2.90	0.0037	0.0036	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028
3.00	0.0027	0.0026	0.0025	0.0024	0.0024	0.0023	0.0022	0.0021	0.0021	0.0020

Survival Analysis: A Practical Approach, Second Edition David Machin, Yin Bun Cheung, Mahesh K.B. Parmar
 ISBN: 0-470-87040-0